Requested Patent: EP1005233A1

Title:

CONSTANT BIT-RATE CODING CONTROL IN A VIDEO CODER BY WAY OF PRE-ANALYSIS OF THE SLICES OF THE PICTURES ;

Abstracted Patent: EP1005233 ;

Publication Date: 2000-05-31 ;

Inventor(s):

MIAN GIAN ANTONIO (IT); BAGNI DANIELE (IT); DE BEI MATTIA (IT); SACCHI MARIA LUISA (IT) ;

Applicant(s): ST MICROELECTRONICS SRL (IT) ;

Application Number: EP19980830599 19981012 ;

Priority Number(s): EP19980830599 19981012 ;

IPC Classification: H04N7/50 ;

Equivalents: JP2000152237

ABSTRACT:

An algorithm based on a pre-analysis of just few slices (GOS) of the current whole picture and/or on a mix of information on the pre-analysis of a slice of the preceding picture and on the actual encoding data of the preceding whole picture is able to implement an efficient constant bit-rate (BR) control with a reduced requisite of buffer memory capacity and a proportionally less costly hardware. The pre-analysis may be carried out by precoding the GOS with a constant reference quantizer or by entropy computation and the local control of the bit-rate is implemented by way of an integrative-proportional controller.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 005 233 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
31.05.2000 Bulletin 2000/22

(51) Int. Cl.7: **H04N 7/50**

(21) Application number: 98830599.1

(22) Date of filing: 12.10.1998

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant:
STMicroelectronics S.r.l.
20041 Agrate Brianza (Milano) (IT)

(72) Inventors:
- Bagni, Daniele
23887 Olgiate Molgora (IT)
- De Bei, Mattia
30019 Sottomarina (IT)
- Mian, Gian Antonio
35100 Padova (IT)
- Sacchi, Maria Luisa
20129 Milano (IT)

(74) Representative:
Pellegri, Alberto et al
c/o Società Italiana Brevetti S.p.A.
Piazza Repubblica, 5
21100 Varese (IT)

(54) Constant bit-rate coding control in a video coder by way of pre-analysis of the slices of the pictures

(57) An algorithm based on a pre-analysis of just few slices (GOS) of the current whole picture and/or on a mix of information on the pre-analysis of a slice of the preceding picture and on the actual encoding data of the preceding whole picture is able to implement an efficient constant bit-rate (BR) control with a reduced requisite of buffer memory capacity and a proportionally less costly hardware.

The pre-analysis may be carried out by precoding the GOS with a constant reference quantizer or by entropy computation and the local control of the bit-rate is implemented by way of an integrative-proportional controller.

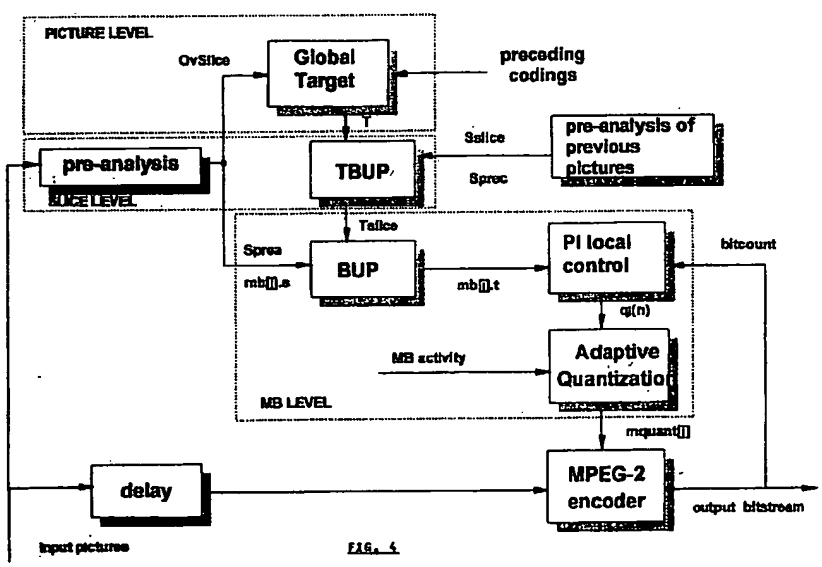


FIG. 1

EP 1 005 233 A1

Description

FIELD OF THE INVENTION

[0001] The present invention relates in general to the processing of digitized pictures and in particular to picture coders, wherein output data streams must have a constant bit-rate (CRB) to satisfy the requisites of the transmission systems and/or of recording systems. The invention is particularly useful in video coders based on the MPEG2 protocol that generates a variable bit-rate (VBR) output stream and that require an output buffer in which restoring the required constant bit-rate by appropriate regulation, but also in coders based on different protocols such as MPEG2, to MPEG1, H263 and H261.

BACKGROUND OF THE INVENTION

[0002] A precise control of the number of bits per Group-Of-Pictures (GOP) is very important when recording the coded video sequence on an appropriate support (for example a CD-ROM) according to the MPEG2 standard. Indeed, it permits simple play back operations such as fast-forward and fast-backward thanks to the fact that the GOP are regularly spaced.

[0003] Only few pictures of the sequence are visualized during these reproducing operations. In particular, only the intra picture (I) are decoded, the only ones allowing a random access because they are independently coded from the rest of the sequence. Therefore, the reproduction consists in searching within the bit-stream and decoding of the first frame (I) of each GOP. Such a search may take place by a jump approximately close to the start of the GOP and then by reading the bit-stream looking for the *start code* of the picture. Alternatively, there is the possibility to insert a limited number of stuffing bits in order to make exactly constant the number of bits per GOP and thus eliminating the searching in the bit-stream.

[0004] Another advantage is the simplification of picture editing. Indeed, it is possible to take a small portion of film, modify it, record it and insert it with precision the same bit-stream interval. Without such a possibility it would be necessary to read and save the entire bit-stream in order to assign the necessary space to the modified sequence.

[0005] The objectives of any CBR system for controlling the bit-rate are:

1) having preliminarly fixed the desired (target) bit-rate and taking into consideration pictures of a certain completely (detail), determining the coding quality and therefore the appropriate parameters for regulating the coding itself in such a way to obtain the predefined bit-rate;
2) maintaining through a control process, a local quality as uniform as possible within the sequence.

[0006] Of course, the ability of accurately control the number of bits produced by each frame has repercussions also upon the quality of the coded sequence. Indeed, when the control system assigns the target bit-rate for the frame to be coded, it also distributes the remaining bits in the GOP among the frames that are still to be coded. If the bits produced by the frame coding are greater in number than the predicted ones, in the GOP will remain less bits than previously envisaged. Thereby, control system will have to diminish the bits at disposal of the successive frames somehow penalizing their quality. These problems occur mainly when there are substantial changes of the picture statistics (and therefore of their complexity). This typically happens upon changes of scene. The two objectives mentioned above are then conflicting.

[0007] In the ensuing description, reference will be made to the following publications on these topics and their pertinent contents recalled as needed:

- [1] International Standard ISO/IEC 13818 (MPEG2), "Information Technology - Generic coding of moving pictures and associated audio", Mar. 1994.
- [2] B. G. Haskell, A. Puri, A. N. Netravali, "Digital video: an introduction to MPEG2", Chapman & Hall, ISBN 0-412-08411-2.
- [3] Test Model Editing Committee, "MPEG2 Video Test Model 5", ISO/IEC JTC1/SC29/WG11 (April 1993).
- [4] G. Keesman, I. Shah, R. Klein-Gunnewiek, "Bit-rate control for MPEG encoders", Signal Processing: Image Communication 6 (1995) pp. 545-560.
- [5] W. Ding, B. Liu, "Rate Control of MPEG video coding and recording by Rate-Quantization modeling", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 1, Feb. 1996.
- [6] S. Bilato, G. Calvagno, G.A. Mian, R. Rinaldo, "Accurate Bit-Rate and quality Control for the MPEG2 video coder", proceedings of ICIP-97 (International Conference on Image Processing), Santa Barbara 1997.

[0008] The article by Keesman [4] describes the finding of a compromise between the two contrasting objectives by

introducing a technique based on defining a so-called Bit Usage Profile which allows to predict how the bits should be distributed within a picture. Such a technique is also described by Mian [6].

[0009] In the majority of video coders there is a control signal which directly influences the regulation of the local quality of a picture. Such signal will be symbolized by *Qj* (where *Q* stands for quantizer). A high value of *Qj* corresponds to a low quality of the image (coarse quantization) while a moderate value corresponds to a higher viewing quality (fine quantization).

[0010] The objective of keeping a uniform quality is then equivalent to having a substantially constant control signal *Qj*, within the picture and possibly throughout the entire sequence of pictures.

[0011] According to the system or model described in the article *"MPEG2 Video Test Model 5"* [3], briefly referred to as TM5, and also according to other CBR models cited in the bibliography, as for example in [5], the control of the bit-rate is typically realized in three phases:

1) estimation of the global target bit-rate (**Global Target** bit allocation), wherein a prediction of the number of bits required for each picture of the sequence is generated. Of course the complexity of the picture and the state of filling of the output data buffer must be considered;
2) determining the reference value *Qj[n]* of the local quantization parameter (**Local Control**) for the n-th macroblock (MB_n) that satisfies previously estimated target bit-rate;
3) calculating the "effective quantization" parameter *mquant[n]* for each macroblock [1], generating it in function of the filling state of the output buffer and of the activity of the macroblock itself (**Adaptive Quantization**).

[0012] While according to the reference model TM5 [3] the latter function is considered a substantial part of the bit-rate control, according to other known systems [4], [5] and [6], as well as in the context of the system of the present invention, is considered a separate problem concerning the conversion of the control signal *Qi[n]* in the regulating parameters of the *mquant[n]* encoder, according to the formula:

$$mquant[n] = Qj[n] \cdot \left(\frac{2 \cdot act[n] + avg\_act}{act[n] + 2 \cdot avg\_act}\right)$$

where *act[n]* is a measure of the detail of the picture based on the variance of the n-th macroblock, while *avg.act* is the average value of *act[n]* on the whole picture. The expression within the brackets is often referred to as the *normalized activity Nact[f]*.

[0013] Therefore, in the ensuing description, the Adaptive Quantization will not be taken into consideration because it is a separate function, and it is not directly involved in the algorithm of the present invention.

[0014] Both the global target estimation and the local control use information on how to distribute the bits among the different pictures of the sequence and within the single pictures. These information may be derived from the knowledge of the previously coded pictures (according to the TM5 [3]), or by means of a modeling of the Rate-Quantization curve (according to the system described in [5]), or derive from a pre-analysis carried out on each picture (as in [4] and [6]).

[0015] In the system described by Mian [6], a pre-analysis is carried out to calculate the entropy of the DCT coefficients, after having approximated the probability density function by way of histograms. By establishing a certain quantizer $Q*$, the entropy *Hn* of the n-th quantized DCT coefficient, is calculated as:

$$Hn = -\sum_{k=-L}^{L} P_{k,n} \cdot (\log_2 P_{k,n})$$

where $P_{k,n}$ indicates the probability to obtain the K-th symbol at the output of the quantization. Therefore the average entropy represents the actual bit-rate *R*:

$$R = H(Q\cdot) = \frac{1}{64} \cdot \sum_{1}^{64} Hn(Q\cdot)$$

[0016] The estimates obtained with the statistical model, when the quantizer $Q*$ for each macroblock is fixed, are

in accordance with the performances of an MPEG2 encoder, wherein $Q*$ represents the average value of the quantizers effectively used [6].

[0017] It should be noted that with this method, only global estimates (that is on a whole picture) of the actual bit-rate R may be obtained, because the histograms require a large number of samples to be statistically valid. Hence, a possible local level estimate of the single n-th macroblock is lost.

[0018] The pre-analysis with pre-coding, according to the method described by Keesman [4], constitutes a first example of reference and comparison and is hereinbelow reported as a basis for subsequently describing the method of the invention and point out the differences.

CONTROL OF THE BIT-RATE WITH PICTURE PRE-ANALYSIS

[0019] The scheme of Fig. 1 highlights the following elements of the functional coding blocks:

a) an MPEG2 **encoder** implementing a hybrid scheme of DCT interframe coding, according to what described in [1] and [2];
b) an **adaptive quantization** block input with the control signal $Qj[n]$ and which outputs the effective quantizer parameter *mquant[n]*;
c) a block computing the estimated global number of bits (**Global Target** bit allocation), using information derived from previously coded pictures and using information derived from a **pre-analysis** of the current picture;
d) a block computing the **Bit Usage Profile** (BUP), using information derived from the above-mentioned pre-analysis;
e) a proportional-integrative **local controller** (briefly **P-I** which obtains the right Qj signal that will produce the required bit-rate.

[0020] This bit-rate control scheme presents some important differences compared to the TM5 model:

1) in the computation of the global target a difference is made between overhead data and the data relative to the coding of the DCT coefficients; (by the expression overhead bits we refer to all the bits not used for coding the DCT coefficients, such as the coding data of motion vectors, slice address, type of macroblock, address of the successive macroblock, type of quantization employed and the like);
2) while in the TM5 model a P controller (proportional) is used, in the system of this invention a P-I controller (Proportional-Integrative) is used;
3) the P-I controller of the bit-rate employs a so-called technique Bit Usage Profile prediction, reducing the number of control actions.

GLOBAL TARGET ALLOCATION

[0021] Let us assume that the complexity of pictures of the same type be constant in time (stationarity property) and to have a number R of bits available for a GOP. The calculation of the target bit-rate is carried out before a true coding reserving a sufficient number of bits for the remaining pictures belonging to the considered GOP.

[0022] Therefore, if $S[n]$ represents the required number of bits to code the n-th picture and if k pictures of the current GOP have already been coded, there will be then Rl bits for coding the remaining N-k pictures of the GOP in question. By calling $T[n]$ the estimated number of bits (target) for a current picture, then;

$$Rl = R - \sum_{n=0}^{k-1} S[n] = \sum_{n=k}^{N-1} T[n] \qquad (1)$$

[0023] Let us assume that the number of bits per picture necessary for the encoding be equal for all the pictures of the same type, in accordance to the above mentioned hypothesis of stationarity; this assumption is reasonable if we consider the coding of pictures of similar contents from the point of view of their complexity, for example by assuming that there are not changes of scene.

[0024] The targets for I, P or B pictures are indicated with Ti, Tp, Tb, respectively. The number of I, P and B pictures inn the GOP that remain coded is indicated with Ni, Np and Nb, respectively. In line with the above assumptions:

$$Rl = Ni \cdot Ti + Np \cdot Tp + Nb \cdot Tb \qquad (2)$$

[0025] The coding of a picture yields two types of data: the bits used to code the DCT coefficients (Huffman coding) and the bits used to code the so-called overhead information (that is all that is not a DCT coefficient). In the following paragraphs we shall conventionally use the symbol "*" to identify the different types of coding for I, P or B frames. The overhead data of a generic picture, denoted by $O*$, are generally independent of the average control signal Qm, which is considered the average of $mquant[n]$ over a picture, while the number of bits for coding the DCT coefficients, denoted by $C*$, depends on the quantization parameters. Therefore:

$$T^* = C^* + O^* \qquad (3)$$

[0026] The number of bits required for the DCT coefficients depends not only on the average control signal Qm but also on the content of the picture, a measure which may be the complexity X^*. Similarly to the TM5 technique, we assume the validity of the following relation for each type of picture:

$$X^* = C^* \cdot Qm \qquad (4)$$

[0027] The complexity $X*$ may be obtained either from previously coded pictures of the same type of the current one or through a pre-analysis, which consists in precoding with a fixed and constant $Qj[n]$. In any case the complexities Xi, Xp and Xb of the different pictures may be assumed known.

[0028] It is assumed to strive to obtain a uniform quality in time by choosing, relatively for the different type of pictures, average values for the control signal with constant ratios:

$$Kip \cdot Qi = Qp \qquad (5)$$

$$Kpb \cdot Qp = Qb$$

where Qi, Qp and Qb are the average values Qm relative to the I, P and B pictures. The constants Kip and Kpb establish the scaling factor of the global complexities used in calculating the global targets. In the original article of Mian, et al., $Kip=1.0$ and $Kpb=1.0$ were selected, while in the TM5 model $Kip=1.0$ and $Kpb=1.4$ are used.

[0029] At this point it is possible to derive the equations for the global targets. Firstly, estimates of the Oi, Op and Ob values are obtained from the pre-analysis. It is also possible to estimate the overhead from the data of the last frame encoded in the same manner of the current one.

[0030] The estimated values for the variables $C*$ are derived by solving a system of equations. For clarity of notation the variable RL is introduced:

$$RL = Rl - (Ni \cdot Oi + Np \cdot Op + Nb \cdot Ob) = Ni \cdot Ci + Np \cdot Cp + Nb \cdot Cb \qquad (6)$$

which denotes the number of bits left for the DCT coefficients without any the overhead data.

[0031] At this stage, there are six equations: the relationship (4) which represents three equations, the two equations (5) and the equation (6), and six unknown variables Ci, Cp, Cb, Qi and Qb. The values of Ni, Np and Nb and the complexity values $X*$ are known from the pre-analysis or taken from previous encodings. The flowing global target equations for the DCT data can be derived by simple algebraical operations:

$$Ci = \frac{Xi \cdot RL}{Ni \cdot Xi + Np \cdot Xp / Kip + Nb \cdot Xb / (Kip \cdot Kpb)} \qquad (7)$$

$$Cp = \frac{Xp \cdot Ci}{Kip \cdot Xi}$$

$$Cb = \frac{Xb \cdot Ci}{Kip \cdot Kpb \cdot Xi}$$

[0032] From equations (7) it may be observed that the global targets distribute the bits among the pictures according to the relative complexities. Abrupt changes of scene are a problem for the computation model of the global targets. In fact, up to here the global complexities are partially obtained from pictures that have been already coded in the same manner. For the B or P pictures, such a prediction does not create a problem even in the case of an abrupt change of

content of the pictures of the sequence. However, for the I pictures there may be some problems because the distance among them may be at 12 frames. The pre-analysis may give more reliable information on the global complexities, in presence of changes of scene. In any case, after a change of scene there is a short period of time in which the human sight is less sensible to the bluffing effect of the pictures. This phenomenon allows the use of global targets lower than what would be really necessary and causes the generation of coding errors in pictures immediately following a change of scene.

[0033] The Mian algorithm [6] employs a similar procedure of Global Target calculation and therefore leads to the same solutions expressed by the equations (7).

LOCAL CONTROL

[0034] The algorithm for estimating the global targets obtains, as explained above, the average value *Qm* of the different *mquant[n]* and the average number of bits per macroblock. If too many bits are used for a portion of the picture the *Qj[n]* of the macroblocks of the remaining portions of the picture must be incremented accordingly in order to obtain a bit-rate of the picture as compliant as possible with the global target, at the expense of the uniformity of the quality.

[0035] For each macroblock, the local control computes the relative control signal *Qj[n]* using information on the macroblocks already encoded and information on the bit usage. The model is thoroughly analyzed in the article [4] which contains an analysis of the stability based on the Theory of Automatic Controls.

[0036] Fig. 2 shows a block diagram of the local control. A unit of time is taken the interval between the encoding of two successive macroblocks.

[0037] Therefore, the n-th instant indicates the encoding instant of the n-th macroblock of the current picture. The variables used in this analysis, functions of time *n*, and indicated in the scheme of Fig. 2 are the following:

- *Qj[n]* is the value of the reference quantization parameter for the n-th macroblock (or briefly MB);
- *s[n]* is the number of bits obtained upon effectively coding of the n-th macroblock;
- *b[n]* is the effective filling state of the output buffer;
- *p[n]* is the estimated number of coding bits for the macroblock, deriving from the information produced at the end of the pre-analysis;
- *t[n]* is the estimated filling state of the buffer (that is the target state);
- *r[n]* is the number of transmitted bits (eliminated from the output buffer) in the time interval of the n-th macroblock. It should be noted that the term *r[n]* appears twice in the scheme: it is subtracted from the number of bits *s[n]* produced to calculate the real filling stage of the buffer *b[n]* and is also subtracted from the estimated number of bits *b[n]* to reach the estimated filling state *t[n]* of the buffer;
- *e[n]* is the error between the estimated and effective filling state of the buffer, that is:

$$e[n]=t[n]-b[n] \qquad (8)$$

[0038] The functional blocks of the scheme of Fig. 2 are:

- PI: the Proportional-Integrative local control block; it selects the value of *Qj[n]* depending on the detected error;
- MPEG ENCODER: the encoding block; it represents the means that transform the selected value *Qj[n]* and the obtained number of bits *s[n]* and also the means of adaptive quantization.
- BUFFER: a data accumulator (memory buffer); its input *x* and output *y* are linked by the formula $y[n] = y[n-1]+x[n]$, which, in terms of the Zeta Transform, becomes:

$$Y(z)=\frac{1}{1-z^{-1}} \cdot X(z)$$

[0039] *E(z)*, *B(z)*, *T(z)* and *P(z)* are the Zeta Transforms of *e[n]*, *b[n]*, *t[n]* and *p[n]*, respectively. The error signal *E(z)* is in the difference-between the buffer content *B(z)* and the desired content *T(z)*. The latter signal is obtained by integrating the bit profile *P(z)* which represents the expected number of bits per macroblock. According to the TM5 model, this number is constant within a picture, that is, it is a hypotized uniform spread of bits among all the macroblocks.

[0040] The local controller, whose architecture is shown in Fig. 3, obtains the *Qj(z)* signal from the error signal *E(z)*. The controller is of the proportionally-integrative (P-I) type, having a proportional path with a *KP* gain and an integrative path with a *KI* gain. From preliminary tests and from the indications present in [4]: *KP=0.01* and *KI=0.001*.

[0041] The following equation links the quantization parameter *Qj(z)* to the *E(z)* error:

$$Qj(z)=\left(KP+KI\frac{1}{1-z^{-1}}\right)\cdot E(z) \tag{9}$$

[0042] Using the inverse transform and by recursively proceeding to the 0 instant, having placed $e[0]=0$:

$$Qj[n]=Qj_{REF}+KP\cdot e[n]+KI\cdot\sum_{u=1}^{n}e(u) \tag{10}$$

where Qj_{REF} is the constant reference quantization parameter for the picture that is used in the pre-analysis and which may be calculated in different ways. Such parameter is the average of the $Qj[n]$ of the preceding picture (regardless of the type), that is:

$$Qj_{REF}=\frac{1}{allMB}\sum_{n=1}^{allMB}Qj[n] \tag{11}$$

[0043] It should be borne in mind that Qm and Qj_{REF} are the average values of $mquant[n]$ and $Qj[n]$, respectively, for the whole picture, and that they are linked by the average value of the normalized activity $Nact[j]$.

[0044] The choice of a proportional-integrative (P-I) controller provides for a regulation without static error while a proportional (P) controller (such as in the TM5 model and also adopted by Mian [6]) is affected by the static error component.

[0045] A compromise is necessary in selecting the controller parameters KI and KP in order to achieve both the stated objectives. If a low gain is set in the control loop, for example a KP close to zero, there is a loss of control effectiveness. In such a case the quality of the picture may be very uniform but the bit-rate will be hardly predictable. An encoder with a relaxed control of the used bits may incur into problems every time there are overflow or underflow situations.

[0046] On the other hand, if a high gain is set, for example a KP close to 1, in order to implement a higher control of the bit-rate, the quality of the picture may become rather inconstant. In order to solve this problem, the Bit Usage Profile is used.

THE BIT USAGE PROFILE

[0047] The objective is to determine in advance the number of bits $t[n]$ that will be necessary for each macroblock, this approach is referred to as Bit Usage Profile. To this end, a pre-analysis is used (see Keesman [4]) according to which a complete encoding of the current picture is done using a constant $Qj[n]$ value. Thence, the number of bits $s[n]$ that each macroblock uses during this precoding process is obtained.

[0048] Similarly to what is done for the global complexity $X*$ of a picture, a local complexity of a macroblock is defined as:

$$x[n]=s[n]\cdot Qj[n] \tag{12}$$

with the following condition:

$$X*=\sum_{allMB}x[n] \tag{13}$$

[0049] The Bit Usage Profile (BUP) of each macroblock is indicated by $t[n]$; it represents the estimated number of bits that each macroblock will use for effective encoding, with the condition that the sum of the various $t[n]$ be equal to the global target $T*$, that is:

$$T^* = \sum_{allMB} t[n] \qquad (14)$$

[0050] By redistributing the *t[n]* bits among the various macroblocks of the picture, depending on their complexity *x[n]*, the following equation is derived:

$$\frac{t[n]}{x[n]} = \frac{T^*}{X^*} = \frac{T^*}{\Sigma x[n]} \qquad (15)$$

[0051] Upon assuming that during precoding the *Qj[n]* is constant **(for example equal to Qj_{REF})**, for the various macroblocks it is obtained:

$$BUP = t[n] = \frac{T^* \cdot x[n]}{\Sigma x[n]} = \frac{T^* \cdot s[n] \cdot Qm}{\Sigma(s[n] \cdot Qm)} = \frac{T^* \cdot s[n] \cdot Qj_{REF}}{Qm \cdot \Sigma s[n]} = T^* \cdot \frac{s[n]}{\Sigma s[n]} \qquad (16)$$

[0052] It should be noted that the TM5 model assumes a uniform bit distribution among the different macroblocks, that is

$$BUP = \frac{T}{\#MB}.$$

[0053] A symbolic level, equation (10), which describes the local control model, remains unchanged and valid, although the value *t[n]* of equation (8) is in reality the BUP described by equation (16).

[0054] The algorithms described in [4] and [6] produce an optimal quality and an accurate control of the bit-rate. However, their main application is substantially restricted to the professional market sector, where processing delays of a certain entity may be tolerated and wherein costly memories may be used eventually in order to resynchronize the video and audio packets. Nevertheless, the pre-analysis inevitably lengthens the processing time of the video stream. On the other hand, the consumer market sector imposes the use of the least amount of memory space in order to limit costs.

OBJECT AND SUMMARY OF THE INVENTION

[0055] Objective of this invention was to find a Constant Bit-Rate (CBR) control method for hybrid DPCM-DCT video coding systems in general (MPEG1, H.263, H.261) and more in particular for coding systems according to the MPEG2 standard, which would imply a reduced memory capacity requisite and therefore reduced hardware costs, without notably loss of the quality of the transferred pictures, as compared to known control systems.

[0056] The method of the invention is able to keep constant the output bit-rate that is produced at the end of the coding process, with a high accuracy, by using a mix of information deriving from the effective encoding of preceding pictures, from the pre-analysis of preceding pictures and from a pre-analysis advantageously executed on a certain group of slices or even on a single slice (GOS) of the current picture; a slice being constituted by at least a whole line of macroblocks of a picture.

[0057] According to preferred embodiments, a GOS pre-analysis may be alternatively carried out:

a) by precoding one or more lines of macroblocks (slice), with a constant reference quantizer;
b) by way of a statistical calculation (with histograms) of the entropy of the DCT coefficients that approximates the average characteristic waveform of the bits necessary to encode the DCT coefficients.

[0058] As compared to a CBR controller of the prior art using a pre-analysis of the whole current picture, the system of the invention provides for an outstandingly accurate control while preserving a good quality of the pictures comparable to those of the known systems, but with a reduced processing delay and a reduced memory requisite for managing pre-analysis data.

[0059] By conducting a pre-analysis on a current GOS rather than on the whole current picture, it is necessary to store information limited to the current GOS, thus reducing the memory requisite, the area occupied on the silicon and the overall cost of the device.

[0060] Even, the time of pre-analysis of a GOS rather than of the whole picture, is much shorter thus facilitating synchronization among the audio and video packets.

[0061] Surprisingly, it has been found that by operating a pre-analysis on a current slice or group of slices of length equal to the width of the picture and in a maximum number of four slices (GOS) at the time, it is possible to retain a picture quality that is indistinguishable from that of the known systems.

[0062] This is obtained by defining a distribution within each slice of the target bits as a reference for the calculation of the quantization parameter Qj at the local control level. This technique is referred to as Target Bit Usage Profile (TBUP) and constitutes an essential aspect of the novel algorithm of the invention as defined in the appended claims.

[0063] According to the method of the invention, not having pre-analysis data of the whole current picture, but just those of a GOS thereof, a mix of the data of the pre-analysis of the GOS of the preceding picture (regardless of its type I, P or B) and of data of the effective encoding of the preceding picture (regardless of its type I, P or B), is effected in order to implement said TBUP technique according to the two alternative pre-analysis methods, mentioned in the above paragraphs a) and b). The requested memory capacity to store the above mentioned data is in any case negligible, being such data related to the number of slices or GOS contained in one whole picture.

[0064] Finally, in order to implement a tight bit-rate control while preserving a substantially uniform quality of the pictures, the Bit Usage Profile (BUP) of each macroblock representing the estimated number of bits that each macroblock will be using for the effective encoding is computed.

[0065] In case of a pre-analysis according to the paragraph a) mode, involving the precoding of a slice or multislice GOS of the current picture, the BUP is obtained by determining in advance the number of bits necessary for each macroblock through the pre-analysis itself.

[0066] In case of a pre-analysis according to the paragraph b) mode, involving the calculation of the entropy on a slice or multislice GOS of the current picture, the BUP cannot be obtained from the pre-analysis itself that is only capable to produce an estimate limited to the GOS being processed and not for each macroblock; however, even for this second embodiment, the BUP may be obtained by determining the number of bits necessary for each macroblock from the information stored during the effective coding of the GOS of the preceding picture.

BRIEF DESCRIPTION OF THE DRAWINGS

[0067]

Figure 1 is a functional block diagram of a CBR control algorithm with pre-analysis of the whole picture, according to the known technique, as discussed above.

Figure 2 is an architectural block diagram of the CBR control function of an MPEG encoder.

Figure 3 is an architectural scheme of a proportional-integrative local controller.

Figure 4 is a functional block diagram of a CBR control algorithm with pre-analysis of a GOS, of the present invention.

Figure 5 is a flow chart of an embodiment of the invention (**K.sl.1**), based on a pre-analysis with precoding of one slice of the current picture to implement the BUP, and based on the use of data resulting from the effective encoding of the preceding picture, of the same type of the current one for implementing the TBUP.

Figure 6 is a flow chart of an embodiment of the invention (**K.sl.3**), based on a pre-analysis with precoding of one slice of the current picture to implement the BUP, and based on the use of data resulting from the pre-analysis with precoding of the preceding picture, of the same type of the current one for implementing the TBUP.

Figure 7 is a flow chart of an embodiment of the invention (**qj_stat**), based on a pre-analysis with entropy computation on a GOS of the current picture for calculating the bits target of a slice within the above mentioned GOS (*Tslice*), and based on using the data resulting from the effective encoding of the preceding picture of the same type for implementing the TBUP, and finally based on a BUP of a simply uniform distribution.

Figure 8 is a flow chart of an embodiment oft he invention (**prof_prec**), based on a pre-analysis with entropy calculation on a GOS of the current picture for calculating the bit target of a slice within the above mentioned GOS (*Tslice*), and based on using data resulting from the effective encoding of preceding picture of the same type for implementing both the TBUP and the BUP.

Figure 9 is a flow chart of an embodiment of the invention (**prof_med**), based on using data resulting from the effective coding of the preceding picture of the same type for implementing both the TBUP and the BUP, and based on the average of data from the effective encoding of the preceding picture and from the pre-analysis and entropy calculation on a GOS of the current picture, for calculating the bit targets of a slice within the GOS (*Tslice*).

DESCRIPTION OF SEVERAL EMBODIMENTS OF THE INVENTION

[0068] By referring to Fig. 4, the ensuing description of the system of the invention follows a scheme to the one used for describing the known system of Fig. 2. The differences between the system of the invention and the known system will be punctually highlighted.

GLOBAL TARGET BIT ALLOCATION

[0069] From the method of the invention the relationships expressed by way of the above equations (3), (4), (5), (6), and (7), still hold.

[0070] According to the invention, by carrying out a pre-analysis on a GOS, at the beginning of the actual encoding, information of the overhead data for the whole picture is not available.

[0071] What is done is a comparison between the number of overhead bits estimated by the pre-analysis on the first GOS multiplied by the number of GOS (which means to consider constant the overhead on each GOS) and the overhead calculated in the pre-analysis of preceding pictures of the same type. The lower of these two values is used for calculating the global target, thus underestimating the number of not DCT bits estimated for the remaining pictures of the GOP and producing a less stringent global target for the current frame.

[0072] For the first P and B pictures of each GOP, excluding the first, the overhead is calculated as the average of the overheads of the preceding GOP so to have a more realistic estimate given the lack of information at the beginning of a new GOP. Since the I picture is not subject to motion compensation, it is characterized by a smaller number of overheads bits than the other types of pictures (P or B) and it is then treated separately by considering the overhead of the corresponding I picture of the preceding GOP.

[0073] The same applies also for the complexity $X*$ (see equation 4), which may no longer be obtained from the pre-analysis, because at the end of the processing of the first GOS there is not enough information on the whole picture. However, the complexity $X*$ is obtained by considering the DCT bits of the pictures of the same type, previously encoded.

BIT USAGE PROFILE

[0074] As already explained above, the BUP permits a tight control of the bit-rate while preserving uniformity of the quality of the pictures. According to the method of the invention, the BUP is obtained by determining in advance the number of bits that will be needed for each macroblock $t[n]$ through a pre-analysis on the slice that is about to be encoded.

[0075] The pre-analysis consists in pre-coding all the macroblocks of the current GOS keeping the quantization parameter Qj constant, similarly to what is done in Keesman [4]. In this way, an estimate of the number of bits used by each macroblock $s[n]$ is produced.

[0076] By considering the restraint that the sum of the various $t[n]$ be equal to the global target of each slice, $Tslice$, the following formula is obtained for calculating the BUP of each macroblock:

$$BUP = t[n] = Tslice \cdot \frac{s[n]}{\Sigma s[n]} \tag{17}$$

where $\Sigma s[n]$ is equal to the sum of the slice bits obtained from the pre-analysis.

[0077] In relation to the equation (16), it is employed a global target for the current slice ($Tslice$) rather than for the whole picture ($T*$). A first way to determine it is to consider the estimated bits uniformly distributed on the different slices, so taking as a reference the uniform distribution of bits among the various macroblocks. This type of solution may deviate from reality when, for instance, a part of the picture requires more bits than another one.

[0078] According to the invention, this drawback is effectively overcome by introducing a Target Bit Usage Profile, as better explained in the ensuing description.

[0079] Finally, it should be noted that for the case of a pre-analysis with calculation of the entropy, as per the approach used by Mian [6], the $s[n]$ may be obtained only by knowing the bits effectively used in encoding the preceding pictures (regardless of the type I, P or B), because the pre-analysis with calculation of histograms generates only global estimates for the GOS rather than for the individual n-th macroblock.

TARGET BIT USAGE PROFILE IN AN APPLICATION WITH THE PRE-ANALYSIS FOLLOWING A KEESMAN APPROACH

[0080] By considering uniform the distribution of the target bit on the slices, we experimentally obtain a decrement of the PSNR as compared to an algorithm with pre-analysis of the whole picture and a worsening of the error on a single macroblock.

[0081] These consequences are effectively remedied by defining a **Target Bit Usage Profile (TBUT)**, which represents an estimate of the target bit profile among the different slices and which is evaluated in the pre-analysis phase or derived from the encoding of preceding pictures.

[0082] At the end of the encoding process or of the pre-analysis of the preceding picture, the number of total bits used by the encoder for each GOS S_{GOS} and the number of total bits for each slice of the GOS being processed, *Sslice*, are stored. By considering for example the j-th GOS (containing at the most 4 slices):

$$S_{GOS} = \sum_{k=1}^{4} Sslice[k] \tag{18}$$

$$Sprec = \sum_{k=1}^{NumMax} S_{GOS}[k]$$

where *NumMax* is the maximum number of GOS in a picture and *Sprec* represents the total number of bits for the whole picture.

[0083] The target for the n-th GOS, representing the Target Bit Usage Profile (TBUP), is given by:

$$TBUP = T_{GOS}[n] = \frac{S_{GOS}[n]}{Sprec} \cdot T^{*} \tag{19}$$

with T_{*} equal to the target of the whole picture, found at the beginning of the encoding process by way of the Target Bit Allocation technique. Of course, the sum of all the T_{GOS} in a picture is equal to T_{*}.

[0084] If the GOS contains only one slice: $Tslice=T_{GOS}$ and $Sslice=S_{GOS}$. In this case the BUP is explicitly given by equation (17) rewritten below as equation (20):

$$BUP = t[n] = Tslice \cdot \frac{s[n]}{\Sigma s[n]} \tag{20}$$

where as usual $\Sigma s[n]$ is equal to the sum of the slice bits obtained by the pre-analysis with precoding of the current slice of the current picture.

[0085] Instead, if the GOS contains more slices (4 at the most), the BUP is more accurate because it tends to become more similar to the BUP of the whole picture (as would be the case if the GOS contains all the slices of a picture), thus equation (17) becomes:

$$BUP = t[n] = T_{GOS} \cdot \frac{s[n]}{\Sigma s[n]} \tag{21}$$

where in this case $\Sigma s[n]$ is equal to the sum of the GOS bits derived from the pre-analysis with precoding of the current GOS of the current picture.

[0086] It should be noted that the values of S_{GOS} and *Sslice* may be obtained in different ways depending on whether a knowledge of the bits calculated during the pre-analysis of the preceding pictures or during their effective encoding is used, in both cases regardless of their type: I, P or B.

TARGET BIT USAGE PROFILE IN AN APPLICATION WITH PRE-ANALYSIS ACCORDING TO MIAN

[0087] At the end of the coding or of the pre-analysis of the preceding picture, the number of total bits S_{GOS} produced by the encoder for each GOS, and the number of total bits *Sslices* per each slice of the GOS being processed are stored with the conditions that, for the j-th GOS (containing at most four slices):

$$S_{GOS} = \sum_{k=1}^{4} Sslice[k] \qquad (22)$$

$$Sprec = \sum_{j=1}^{NumMax} S_{GOS}[j]$$

where *NumMax* is the maximum number of GOS in a picture, while *Sprec* represents the total number of bits used for the preceding whole picture, either encoded or pre-analyzed.
[0088] The target for the n-th GOS, representing the Target Bit Usage Profile (TBUP), is given by:

$$TBUP = T_{GOS}[n] = \frac{S_{GOS}[n]}{Sprec} \cdot T^{*} \qquad (23)$$

where T^{*} is equal to the target for the whole picture, found at the beginning of the encoding process by way of the Global Target Bit Allocation; of course, the sum of all the T_{GOS} in a picture must be equal to T^{*}.
[0089] In calculating *Tslice*, only data derived from pre-analysis with entropy calculation of the current GOS of the current picture, that is:

$$Tslice[k] = T_{GOS}[j] \cdot \frac{Eslice[j][k]}{E_{GOS}[j]} \qquad (24)$$

where E_{GOS} represents the number of total bits of the j-th GOS being processed, estimated by way of the entropy, while *Eslice[j][k* is the number of estimated bits of the single k-th slice, with the condition that, for the j-th GOS, it is:

$$E_{GOS}[j] = \sum_{k=1}^{4} Eslice[j][k] \qquad (25)$$

[0090] The above equation (24) is not necessary in a pre-analysis conducted according to Keesman, because if the GOS contains only one slice, *Tslice* is calculated with equation (20) otherwise T_{GOS} is calculated with equation (21).
[0091] Finally equation (17) which calculates the BUP may be transformed in the following expression:

$$BUP = t[n] = Tslice \cdot \frac{s[n]}{\Sigma s[n]} \qquad (26)$$

where, in this case, *s[n]* and *Σs[n]* are obtained from the effective coding of the corresponding GOS of the preceding picture, because information on the current picture is not available.
[0092] Fundamentally, equations (20) (Keesman) and (26) (Mian) are alike, the only difference being that *s[n]* and *Σs[n]* are derived from pre-analysis with precoding (equation (20)) of the current GOS (containing only one slice) of the current picture, otherwise from the effective encoding data of the corresponding GOS (including more than one slice) of the preceding picture (equation 26).

REQUIRED MEMORY SIZE FOR PRE-ANALYSIS DATA

[0093] As already mentioned, the fact of carrying out a pre-analysis only on the GOS involves a large reduction of the memory that is necessary for storing the data. Indeed, by assuming to use a 32 bit precision arithmetic architecture (of the "float" type according to the computer language C), the memory occupation in case of pre-analysis on a four slice GOS typically may be 4 x 32 bit 45 macroblocks per slice, that is, a total of 5,760 bits for any significant information (DCT bits, overhead bits, total bits). By contrast, in case of pre-analysis conducted on a whole picture, according to a prior art system, there the memory occupation would be 32 bit x 1,620 macroblocks, that is 51,840 bits for significant information. This implies a saving of 46,080 bits, or of about 88%. If the GOS contains only one slice, the memory saving rises to 97%.

[0094] The memory capacity necessary to store the *Sslice* information, deriving from a pre-analysis of effective coding (of the preceding picture), is negligible, being it equal to the only number of slices in a picture (there are 36 slices in a PAL picture of 576 active lines), that is 32 bit x 36 slices = 1,152 bits.

LOCAL CONTROL

[0095] The controller used for a local control is of the integrative-proportional type. Once the combined processings according to the known **Bit Usage Profile** technique (alternatively according to one of the equations (20), (21) and (26)) and according to the technique of the present invention that introduces the **Target Bit Usage Profile** (alternatively according to one of equations (19), (23) and eventually also (24) for a Mian compliant version), have been completed, in order to find the target number of bits for each macroblock *t[n]*, it is possible to determine the *Qj* control parameters at the local level (equation (10)), which may then be used, through an adaptive quantization, for calculating the real quantization parameter *mquant*.

[0096] Symbolically, equation (10), which describes the local control model, remains valid and unchanged, with the only difference that only the *t[n]* of equation (8) is in reality the BUP that is alternatively determined by the equations (20), (21) or (26).

[0097] Of course, for calculating the local error of each first macroblock of a slice, it is necessary to store in the memory the datum corresponding to the target number of bits found for the last macroblock of the preceding slice. In this way, it is avoided that the slices be not considered as separate items so that the transition from one to the other may result visible in the encoded picture.

EXPERIMENTAL MEASUREMENTS

[0098] This section reports the experimental results of different implementations of the method of the invention of GOS pre-analysis. As samples of video sequences have been used the so-called Calendar, Stefan and Flowers sequences, coded at 4 Mbps (Megabit per second) and at 10 Mbps.

[0099] It is reminded that not having available pre-analysis data of a whole current picture but only those of a GOS (which may include four slices at the most), a mix is done of the pre-analysis data of the preceding picture (regardless of its type I, P or B) and/or of the data deriving from the effective coding of the preceding picture (regardless of its type I, P or B), in applying the equations (19) and (23) (TBUP), according to an embodiment of the method of the invention compliant with the known method of Keesman and/or of Mian.

[0100] However, to apply equation (26) (BUP), in a Mian compliant embodiment only data originating from coding the preceding picture may be used. By contrast, the equations (20) and (21), according to an embodiment compliant with the method of Keesman, require data derived from the pre-analysis of the current GOS of the current picture.

[0101] Finally, according to an embodiment compliant with the method of Mian, to implement equation (24) for *Tslice*, only data originating from the pre-analysis of the coding of the current picture may be used.

PRE-ANALYSIS WITH PRECODING OF ONLY ONE SLICE

[0102] With regard to the pre-analysis with a precoding, a pre-analysis on a GOS with only one slice has been simulated. The embodiments of the algorithm tested were the following:

1) **K.sl.1**: *Kpb=1.0* the TPUB (S_{GOS} and *Sprec* of equation (19)) is derived from the coding bits of the immediately precedent picture of the same type (I, P, B), as shown in Fig. 5.
2) **K.sl.2**: *Kpb=1.0* and the TPUB (S_{GOS} and *Sprec* of equation (19)) derived from the bits of the pre-analysis the immediately preceding picture, regardless of its type.
3) **K.sl.3**: *Kpb=1.0* and the TPUB (S_{GOS} and *Sprec* of equation (19)) derived from the pre-analysis bits of the immediately preceding picture of the same type (I, P or B), as shown in Fig. 6.

4) **K.sl.4**: $Kpb=1.0$ and the TPUB (S_{GOS} and $Sprec$ of equation (19)) derived from the pre-analysis bits of the immediately preceding picture, regardless of its type if P or B, if of type I the pictures were separately considered by referring to the picture I of the preceding GOP.

5) **K.sl.5**: same as **K.sl.3**, the difference being that during the Target Bit Allocation phase to the I pictures is assigned the estimated number of bits incremented by 20% while the one to the P pictures is incremented by 5%, to produce a somewhat less accurate bit-rate control but an enhanced uniformity of the quality of the pictures.

6) **K.sl.6**: same as **K.sl.4**, however $Kpb = 1.4$ instead of $Kpb = 1.0$.

7) **K.sl.7**: same as **K.sl.3**, however $Kpb = 1.4$ instead of $Kpb = 1.0$.

[0103] For all these test embodiments, the pre-analysis on the current GOS of the current picture is only used for the applying equation (20), that is to estimate $s[n]$ and $\Sigma s[n]$.

[0104] Table 1 compares the performance of the slice algorithms of the invention with that of the method of Keesman [4] and of the TM5 model [3], for the Calendar sequence coded at 4 Mbps. The figures of interest are:

- the average percentage error for the GOP (Err. GOP), calculated as the difference between the Target Bits of a single GOP and the effective number of bits used in encoding, averaged among all the GOPs of the sequence;
- average percentage error among the target bits and those effectively used, for each picture; averaged for the whole sequence (Err. SEQ);
- the picture PSNR averaged on the whole sequence;
- the PSNR of each single macroblock, averaged on all the pictures of the sequence. This measure takes into a greater consideration the artifacts at the macroblock level, which would be inevitably hidden by the intrinsic averaging of the PSNR parameter, that traditionally is calculated on the whole picture;
- the *mquant* average for all the macroblocks of the sequence;
- the mean deviation of the *mquant* (for the entire sequence) in respect to the above average value.

[0105] It may be noted that the TM5 model gives the highest PSNR, though to the detriment of a less accurate control of the bit-rate (average error of 1.61% on the GOP and of 0.19% on the sequence). The Keesman algorithm [4], with the pre-analysis of the whole current picture, provides a very accurate control (average error of 0.0028% on the GOP and of 0.00006% on the sequence), better by least 3 to 4 orders of magnitude than the TM5. By contrast, the different implementations of the method of the invention of slice pre-analysis, loose at most one order of magnitude in terms of the precision over a sequence but ensure a definitively more accurate control than the TM5 model. Differences of the average PSNR among macroblocks are even smaller. Table 1 reports these results.

Table 1

Algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs	Mquant average on sequence	Mean deviation on sequence
TM5	1.61	0.19	27.88	30.78	32.0	5.34
Keesman	0.0028	0.00006	27.03	29.99	32.7	2.0
K.sl.1	0.01	0.0009	26.3	29.7	37.1	3.55
K.sl.2	0.004	0.00008	26.97	30.0	34.1	2.7
K.sl.3	**0.003**	**0.0003**	**26.95**	**29.96**	**33.6**	**2.4**
K.sl.4	0.003	0.0002	26.92	29.94	33.9	2.5
K.sl.5	0.007	0.06	27.46	30.47	32.8	3.8
K.sl.6	0.01	0.00018	27.64	30.73	34.8	2.5
K.sl.7	0.02	0.0008	27.8	30.8	34.0	6.7

[0106] With regard to the measures of the quality of the pictures in terms of PSNR, the Keesman algorithm looses about 0.8 dB as compared to the TM5 model. However, such a degradation is considered negligible according to subjective tests (the PSNR is an objective measure that becomes well correlated to visual evaluations starting from differences above 0.8-1.0 dB). The PSNR differences between the various test embodiments of the method of the invention, based on a slice pre-analysis, and the method of Keesman are much smaller than said range and often the comparison

is in favor of the slice pre-analysis method of the invention.

[0107] The **K.sl.3** algorithm may be considered the best, and for such a preferred implementation the results obtained for the Stefan and Flowers sequences (coded at 4 Mbps) are also reported in the following table.

Table 2

Sequence/algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs	Mquant average on sequence	Mean deviation on sequence
Stefan/TM5	2.02	0.12	30.0	33.4	29.9	7.6
Stefan/Keesman	0.006	0.0006	29.6	33.1	29.7	5.9
Stefan/K.sl.3	0.01	0.00008	29.3	32.9	33.0	7.7
Flowers/TM5	1.03	0.007	28.33	32.72	32.1	6.25
Flowers/Keesman	0.0014	0.0003	27.59	31.96	32.49	2.49
Flowers/K.sl.3	0.002	0.0005	27.39	31.84	33.2	2.7

[0108] As far as quantization is concerned, the above tables report also the average value of the *mquant* of all the macroblocks of the sequences, as an index of a spatial quality; a high value meaning a strong quantization and a consequent blurring of the picture, while a low value of *maquant* meaning a finer quantization and thereby a minimum loss of picture content.

[0109] From this point of view, it is evident that the TM5 model, having a less accurate bit-rate control, has average values consistently lower than the other algorithms with precoding. Nevertheless, the mean deviation of TM5 is higher (with the only exception for the Stefan sequence) than in the algorithms with precoding. This means that the local quality of the algorithm with precoding is more uniform over the whole picture.

[0110] Such an effect may compensate the slight increment of the average *mquant*. Indeed, it may be better to watch pictures of uniform quality, even if slightly more quantized, rather than pictures with zones of significant quantization differences (of nonuniform quality). The optimal performances of the **K.sl.3** version of the algorithm of the invention compared to those obtained with precoding of the whole picture, should be noted. Relatively to mean deviation, even with a limited information derived form the pre-analysis of a single slice, a mean deviation very close to the one that may be obtained with the pre-analysis of the whole picture is obtained.

[0111] Finally, the following table reports these measures on the sequences coded at 10 Mpbs.

Table 3

Sequence/algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs	Mquant average on sequence	Mean deviation on sequence
Calendar/TM5	0.54	0.1	33.2	35.6	13.4	2.22
Calendar/Keesman	0.0003	0.00009	32.36	34.88	13.9	1.2
Calendar/K.ls.3	0.0006	0.00007	32.36	34.95	14.0	1.3
Stefan/TM5	0.3	0.05	35.73	37.84	11.69	2.53
Stefan/Keesman	0.0003	0.0001	35.74	37.8	11.5	2.0
Stefan (K.sl.3)	0.001	0.00008	35.42	37.62	12.6	2.9
Flowers (TM5)	0.2	0.03	33.7	37.27	13.43	2.43

Table 3 (continued)

Sequence/algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs	Mquant average on sequence	Mean deviation on sequence
Flowers (K. orig.)	0.002	0.00006	33.45	36.79	13.4	1.5
Flowers (K.sl.3)	0.002	0.0001	33.36	36.8	13.6	1.6

[0112] By increasing the bit-rate, the average values of PSNR and *maquant* become very similar for all the three methods compared; namely: TM5, precoding of a complete picture and precoding of a single slice (invention). The accuracy of the control is definitely better in precoding systems.

THE PROBLEM OF SCENE CHANGES

[0113] In order to test the behavior of the different methods during changes of scene, some tests simulations were carried out on a sequence (Puzzle) of 50 pictures coded at 4 Mbps, artificially generated by copying frames belonging to different sequences. In particular, the pictures 1 to 7 were taken from the Calendar sequence, 8 to 21 from a Tennis Table sequence, 22 to 23 from Edit, 34 to 38 from Voitur and 39 to 50 from Popple.

Table 4

Algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs	Mquant average on sequence	Mean deviation on sequence
TM5	3.85	0.35	29.6	31.6	30.16	8.89
Keesman	0.002	0.00005	29.14	31.2	30.41	8.42
K.sl.3	0.01	0.0008	28.65	31.0	33.79	8.9

[0114] From the results of Table 4, it may be noted that said how the bit-rate control is more accurate for the two pre-analysis algorithms as compared to the TM5 one. The slight increase of the error and the PSNR with the pre-analysis algorithm of the invention involving only a single slice (K.sl.3), in respect to the one with predecoding of the whole picture (Keesman), is due to the fact that despite the exploitation of information derived from pre-analysis of the current slice of the current picture, some parameters, such as the complexity and the overhead, are inevitably calculated by referring to past pictures. Instead, the complete picture pre-analysis algorithm of the prior art calculates all the needed quantities for the current picture before its actual encoding, thus better determining an intervening change of scene.

[0115] To improve even further the control, in a method according to the present invention, it is possible to use additional information deriving it from the motion estimation block, so as to improve the ability to recognizing changes of scene. Accordingly, the picture corresponding to a change of scene may be coded as an I-picture, obviously if the bit-rate control system has a sufficient number of bits available. The capacity to carry out a replay with fast-forward or with fast-rewind, based on a fixed dimension of GOP (e.g. 12 pictures), is not lost anyway because there will always be at least one I-picture every 12 pictures (in a GOP of a fixed size).

PRE-ANALYSIS BASED ON THE ENTROPY CALCULATION ON MORE THAN ONE SLICE

[0116] In the case of an alternative embodiment of the method of the invention, wherein the pre-analysis is carried out by way of a calculation of the entropy, a pre-analysis on a GOS with 1, 2, 3 and/or 4 slices was hypothized. The results relate to the following variants tested with a 4-slice GOS.

1) ***QJ_stat***: the reference Qj of the whole picture is assumed computed from the pre-analysis with entropy calculation on the first GOS of the current frame. The sum of the estimated bits per slice (E_{GOS} and $Eslice$ of equation (24)) is computed by way of a pre-analysis of the GOS of the current picture. Instead of using the BUP of equation (26), a uniform bit distribution within the single slice is assumed. The TBUP of equation (23) exploits the information S_{GOS} and $Sprec$ relative to the effective encoding of the preceding picture, as illustrated in Fig. 7.

2) **Prof_prec**: the reference *Qj* is calculated as in *Qj_stat*. The bit sum per slice (E_{GOS} and *Eslice* of equation (24)), is calculated by way of a pre-analysis of the GOS of the current picture. The BUP (*s[n]* and $\Sigma s[n]$ of equation (26)) is derived from the effective bits of the previously coded picture. The TBUP of equation (23) exploits the information S_{GOS} and *Sprec* relative to the effective encoding of the preceding picture, as illustrated in Fig. 8.

3) **Prof_medio**: the reference *Qj* is calculated as in *Qj_stat*. The bit sum per slice is obtained by averaging two values, the first value is obtained by exploiting the pre-analysis of the current GOS of the current picture, the second value is obtained from the effective bits of the previously coded picture. The BUP is calculated as in Prof_prec. Finally, a $Kpb = 1.4$ is used instead of $Kpb = 1$ as used in the two preceding test embodiments. The algorithm is illustrated in Fig. 9.

[0117] By coding the sequences at 4 Mbps, the results shown in the following table are obtained. It appears evident a close comparison between the TM5 model and the control methods that use a pre-analysis with calculation of the entropy either carried out on the whole current picture, as in the known Keesman approach, or on single GOS according to the present invention.

Table 5

Sequence/algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs
Calendar/TM5	1.6126	0.1337	27.88	30.78
Calendar/Qj_stat	0.0276	0.00803	26.95	29.61
Calendar/Prof_prec	0.0736	0.00700	26.21	28.94
Calendar/Prof_medio	0.0509	0.00715	27.08	29.94
Flowers/TM5	1.6568	0.20034	28.36	32.72
Flowers/Qj_stat	0.0706	0.00258	27.63	31.57
Flowers/Prof_prec	0.2142	0.00404	26.70	31.04
Flowers/Prof_medio	0.2712	0.00567	27.68	31.91

[0118] It should be pointed out that the performances of the algorithm of the invention with pre-analysis carried out by way of the calculation of the entropy, are approximately midway between those of the TM5 model and those of with precoding. This is so because the entropy is just an approximation measure of the effective bit-rate. The pre-analysis with a precoding is more reliable in calculating the bit-rate. In any case, the control through a calculation of the entropy is simpler to implement, because it is sufficient to produce the histograms of the DCT coefficients rather than performing a precoding.

[0119] By coding the sequences at 10 Mbps, the following results are obtained.

Table 6

Sequence/algorithm	Error GOP (%)	Error SEQ (%)	PSNR average on frames	PSNR average on MBs
Calendar/TM5	0.67	0.12333	32.51	34.99
Calendar/Qj_stat	0.0288	0.00726	31.86	34.16
Calendar/Prof_prec	0.0102	0.00186	30.49	33.47
Calendar/Prof_medio	0.0045	0.00313	31.76	34.38
Flowers/TM5	0.9623	0.17519	33.05	36.72
Flowers/Qj_stat	0.0142	0.00183	32.9	35.94
Flowers/Prof_prec	0.0058	0.00041	32.05	35.53
Flowers/Prof_medio	0.0071	0.00023	32.58	36.09

[0120] In this case, it may be noticed that by increasing the bit-rate, the average PSNR values are very similar for all the compared methods, however the accuracy of the control is definitively better for pre-analysis systems.

[0121] In particular, by considering the results at 4 Mbps and at 10 Mbps, the **Prof_medio** embodiment of the algorithm of the invention shows the best performance overall.

Claims

1. A method of controlling constancy of the bit-rate at a certain value (CBR), depending on optimization criteria in transferring a data stream of encoded picture sequences, comprising the steps of

   a) estimating (Global Target Bit Allocation) the number of bits necessary for encoding a current picture and the successive pictures belonging to the same group of pictures or GOP of a sequence, based on a target number of ($T*$) coding bits for each picture of a sequence, in function of encoding data ($O*$, $X*$, $Q*$) of pictures preceding the current one and/or on the results of a pre-analysis;
   b) determining (Local Control) the reference value (*Qj[n]*) of the quantization parameter at a local level for the n-th macroblock of data, that complies with the limit established by said estimated number of bits for the n-th macroblock (*t[n]*);
   c) calculating (Adaptive Quantization) of an effective quantization parameter (*maquant[n]*) for each macroblock in function of the filling state of an output buffer of said data stream of coded data and of an activity parameter of the macroblock (Nact);
   d) calculating a distribution profile of target bits over a picture (Bit Usage Profile or BUP), by way of a pre-analysis of a certain number of lines of macroblocks or slices (GOS) of the current picture, distributing the allocated bits in function of the local complexity (*x[n]*) of the single n-th macroblock and of the number of bits (*s[n]*) effectively used in said pre-analysis;
   e) calculating a distribution profile of target bits (Target Bit Usage Profile or TBUP) over said plurality of lines of macroblock or slices (GOS) given by the ratio between among the estimated number of bits required for the encoding (S_{GOS}, *Sprec*) obtained by way of a pre-analysis of the current GOS of the current picture or from the effective coding of the corresponding GOS of the preceding picture, multiplied by said target number of bit ($T*$) for of the whole picture;
   said pre-analysis of the current GOS being carried out by precoding the GOS using a constant reference quantization parameter.

2. The method according to claim 1, characterized in that in step d) (Bit Usage Profile or BUP), the effective encoding of said certain number of whole lines of macroblocks or slices (GOS) of the preceding picture is used for distributing the allocated bits in function of the local complexity of the n-th single macroblock (*x[n]*) and of the number of bits effectively used during said effective coding (*s[n]*), and in that after step e) the method includes the following step

   f) intermediately computing the distribution profile of the target bit among the different lines of macroblocks or slices as the ratio between entropy values (E_{GOS}, *Eslice*), derived from a pre-analysis of the current GOS of the current picture, multiplied by said target number of bit usage profile (TBUP);
   said pre-analysis of the current GOS being carried out by way of the calculation of the entropy on the histograms of the DCT coefficients of the whole GOS.

3. A video encoder comprising means for controlling the constancy of the bit-rate at a certain value (CBR), depending on optimization criteria in transferring a data stream of encoded picture sequences, said means comprising an output buffer, an integrative-proportional controller input with the state of filling of said output buffer and outputting a bit-stream which is transmitted through a transmission channel and a control signal for varying the value of the reference quantization parameter for an n-th macroblock of data by said encoder, characterized in that said controller uses estimated data of a certain number of bits (*t[n]*) necessary for encoding a current picture based on a target number of bits ($T*$) for each picture of a sequence determined function of pre-analysis data of a certain number of lines of macroblocks or slices (GOS) of the current picture (*s[n]*, *Eslice*, E_{GOS}) and of the data of an effective coding or of pre-analysis data of a certain number of lines of macroblocks or slices (GOS) of the preceding picture (*Sprec*, S_{GOS}), such a pre-analysis being undertaken by precoding said number of slices (GOS) using a constant reference quantization parameter or by a statistical computation of the entropy of the discrete cosine transform (DCT) coefficients of the macroblocks of said lines of macroblocks or slices.

4. The decoder according to claim 3, characterized in that it comprises a memory for storing said pre-analysis data of said number of lines of macroblocks or slices (GOS) of the current picture (*s[n]*, *Eslice*, E_{GOS}) and the data of an

effective coding or of a pre-analysis of a certain number of lines of macroblocks or slices of the preceding picture (*Sprec*, S_{GOS}), said pre-analysis being carried out either through a precoding of said number of slices (GOS) using a constant reference quantization parameter or by estimating the entropy of the discrete cosine transform (DCT) coefficients of the macroblocks of said lines or slices (GOS).

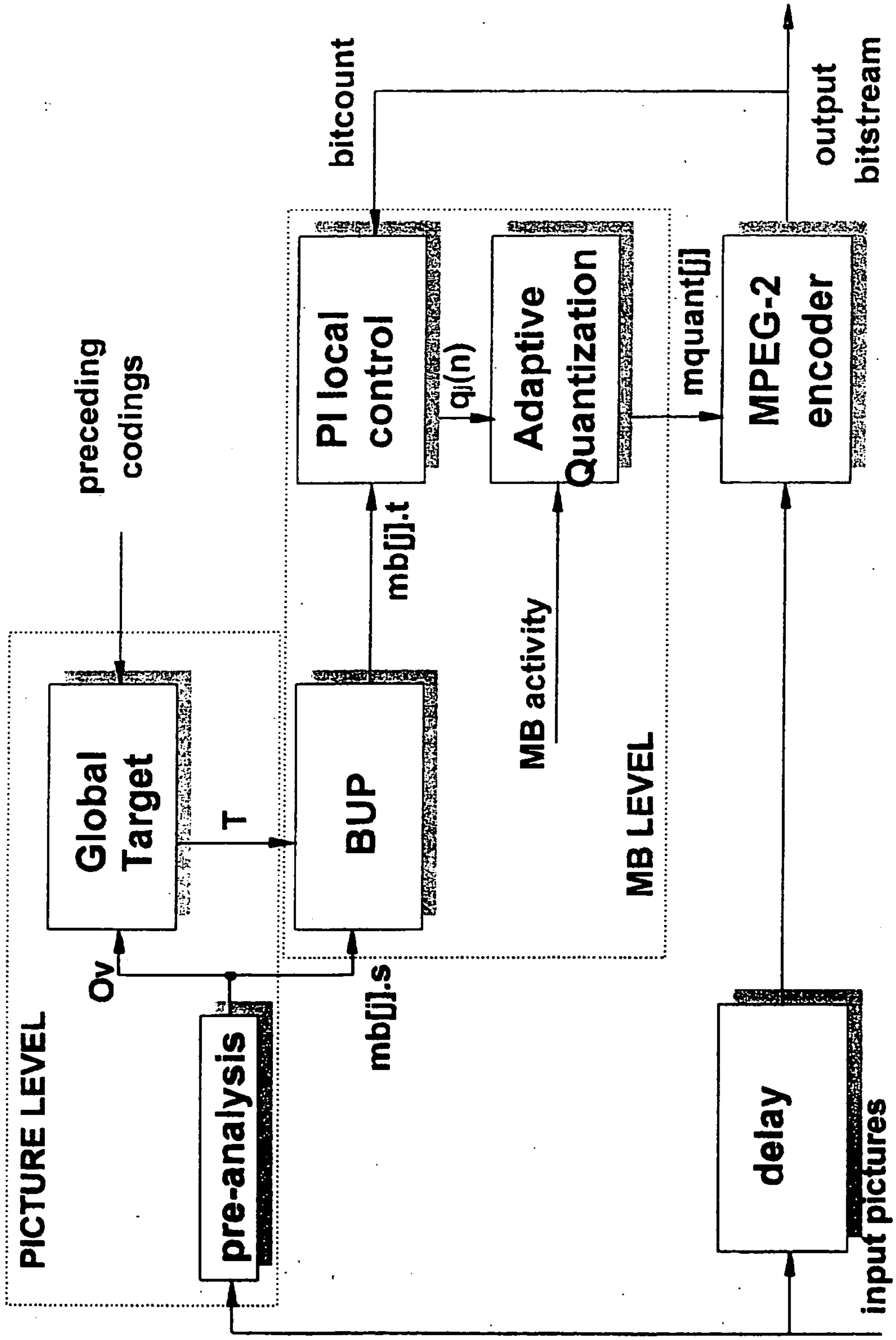

PICTURE LEVEL
pre-analysis
Ov
Global Target
T
preceding codings
BUP
mb[j].s
mb[j].t
PI local control
qj(n)
Adaptive Quantization
MB activity
MB LEVEL
bitcount
mquant[j]
MPEG-2 encoder
output bitstream
delay
input pictures

FIG. 1

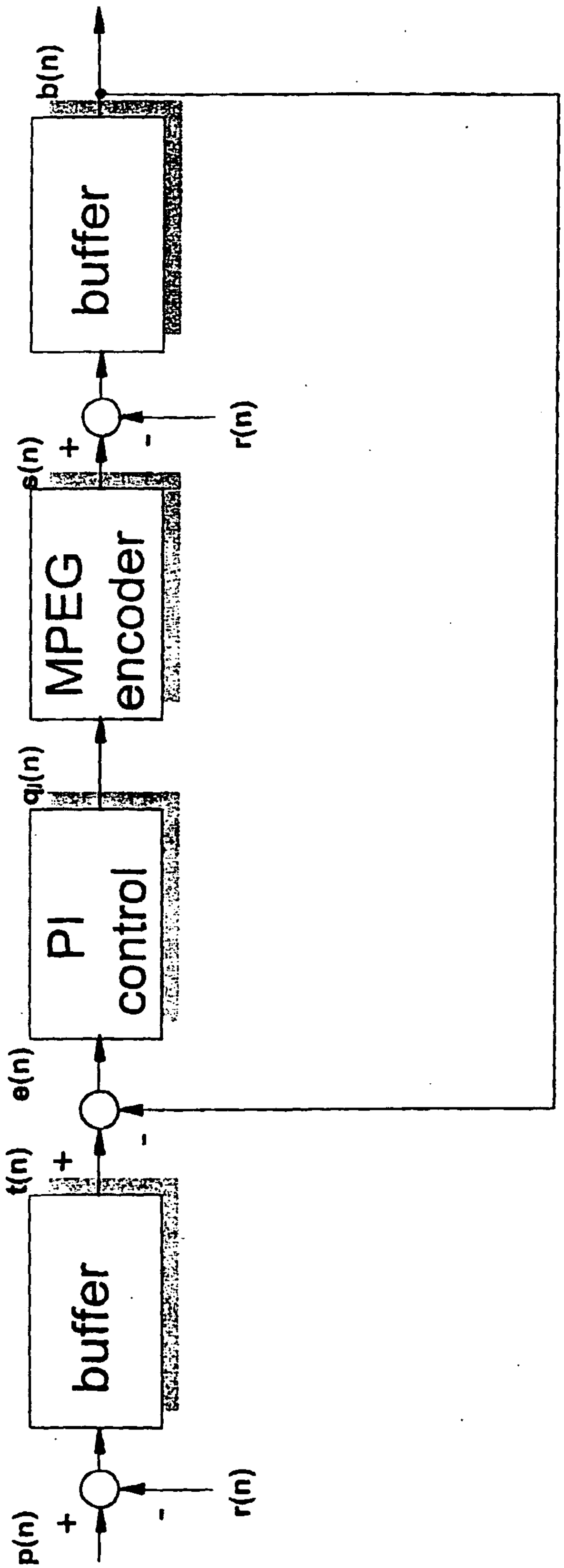


FIG. 2

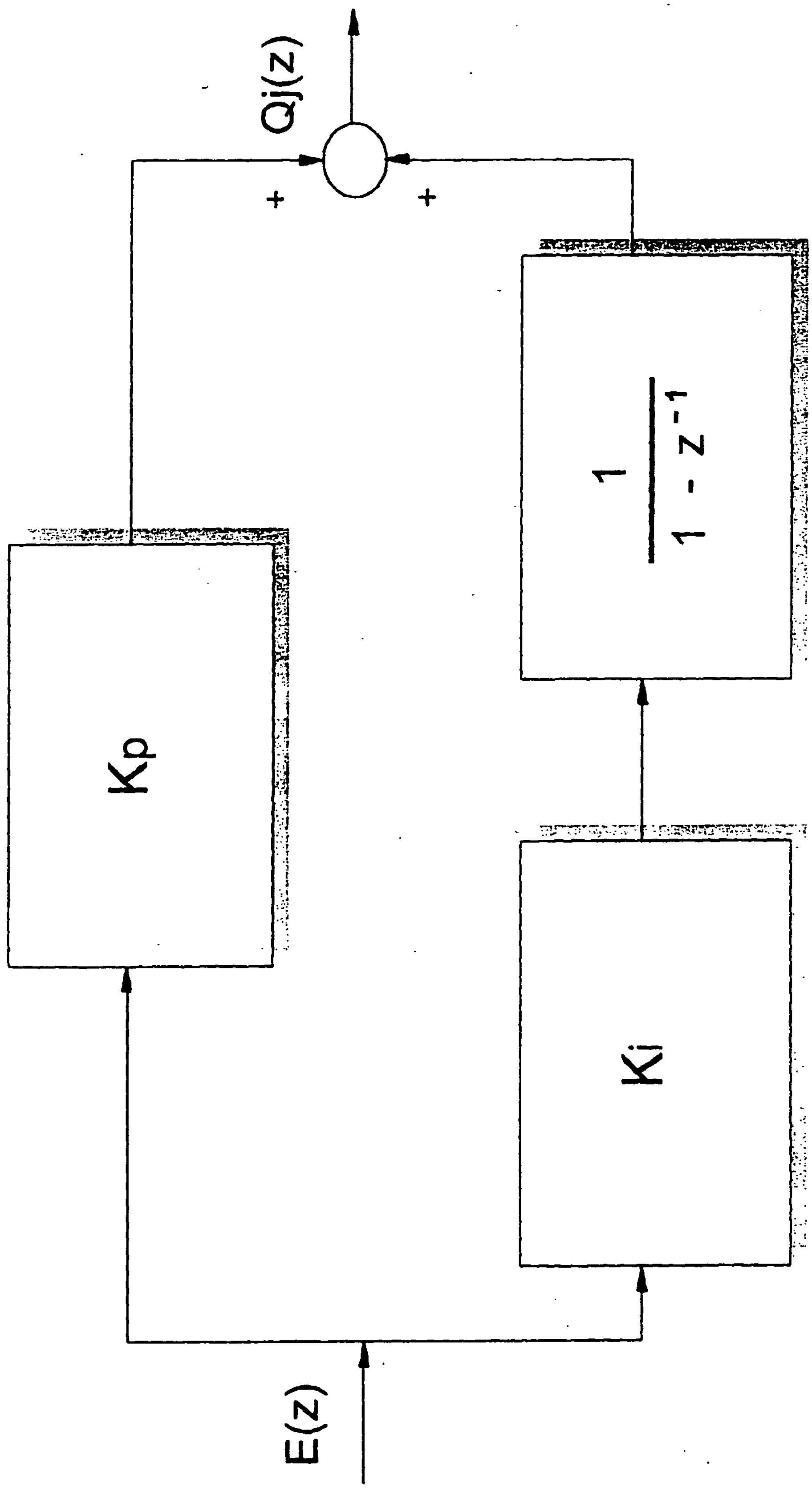


FIG. 3

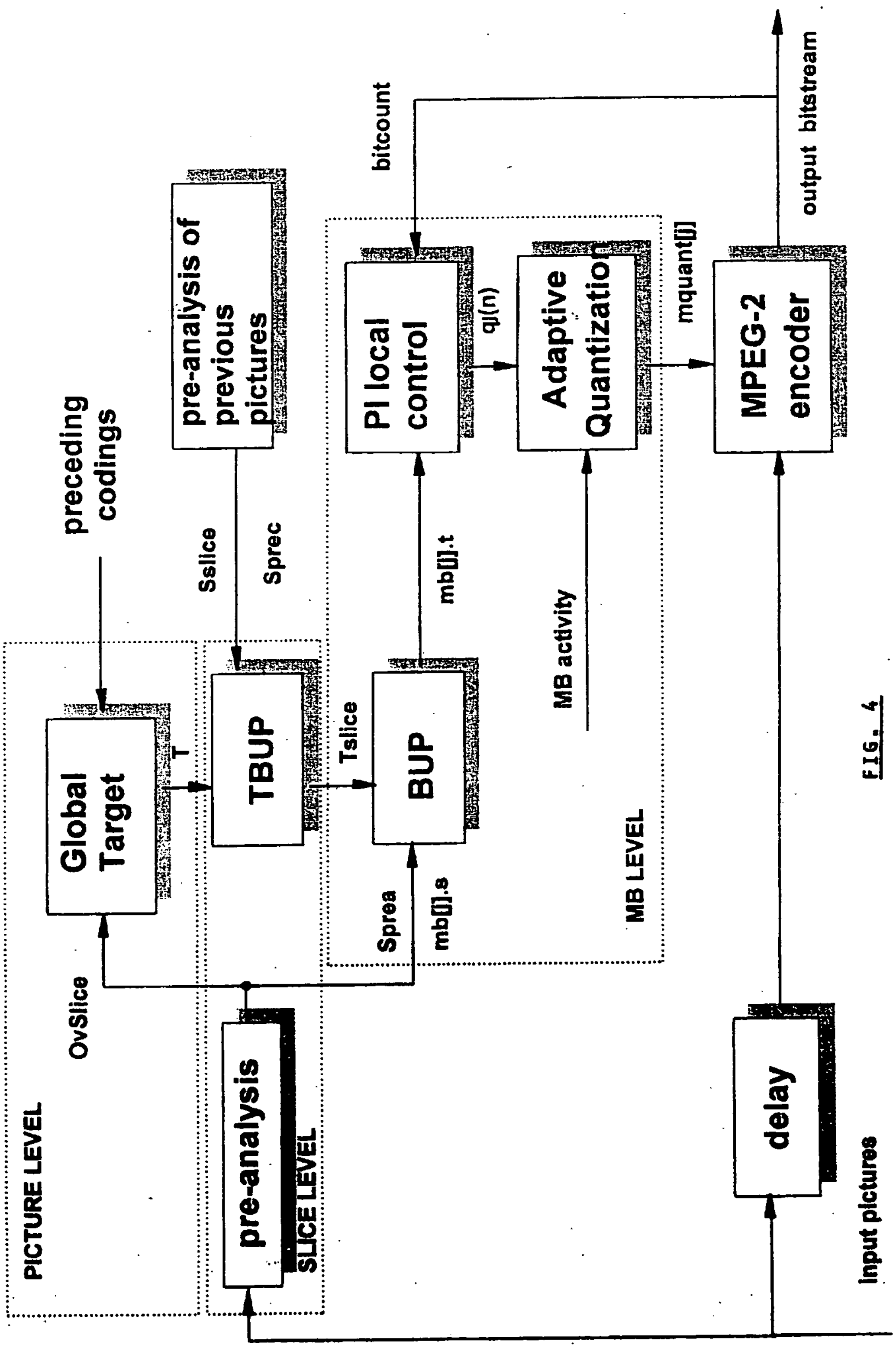

preceding codings
pre-analysis of previous pictures
bitcount
output bitstream
PI local control
Adaptive Quantization
MPEG-2 encoder
qj(n)
mquant[j]
Sslice
Sprec
mb[j].t
MB activity
Global Target
T
TBUP
Tslice
BUP
Sprea
mb[j].s
MB LEVEL
OvSlice
PICTURE LEVEL
pre-analysis
SLICE LEVEL
delay
Input pictures

FIG. 4

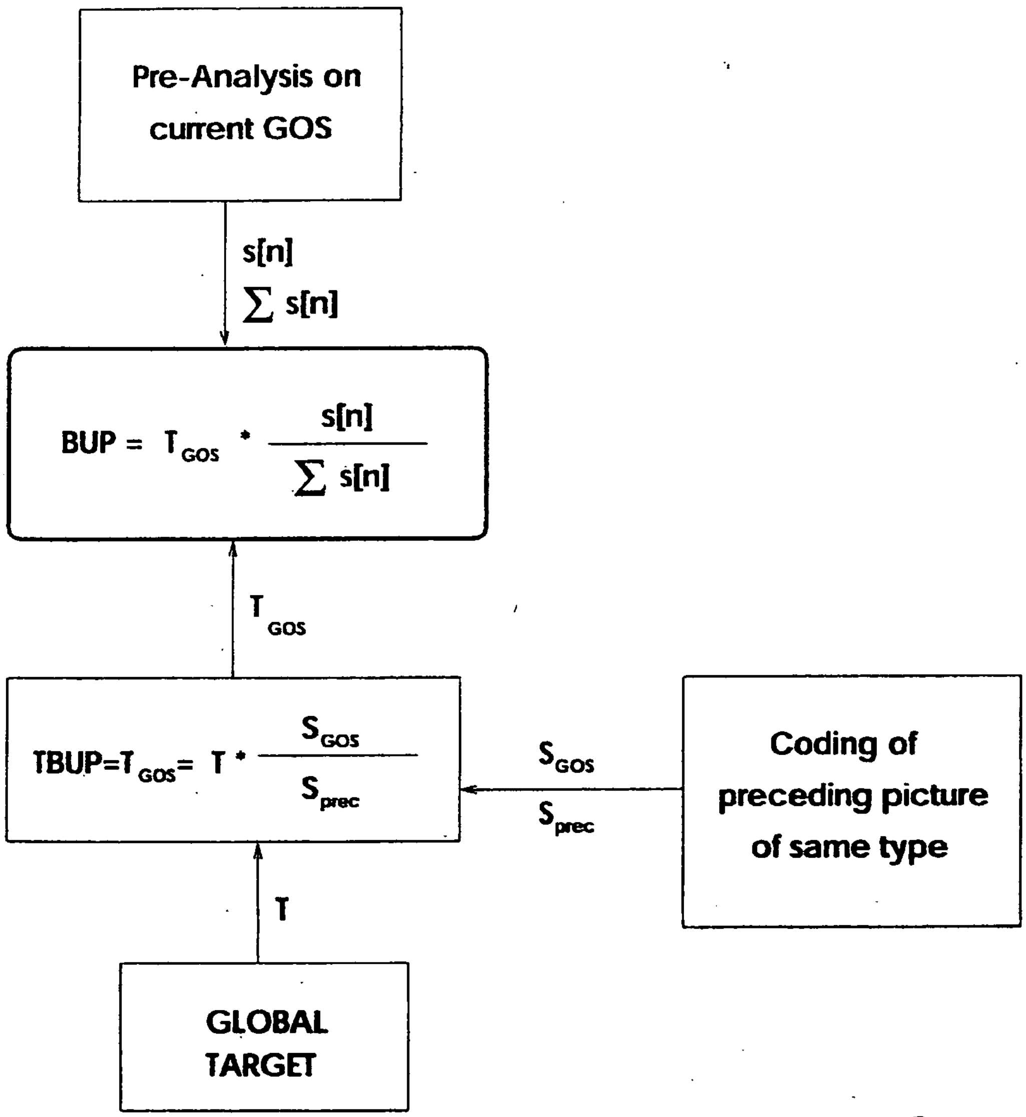
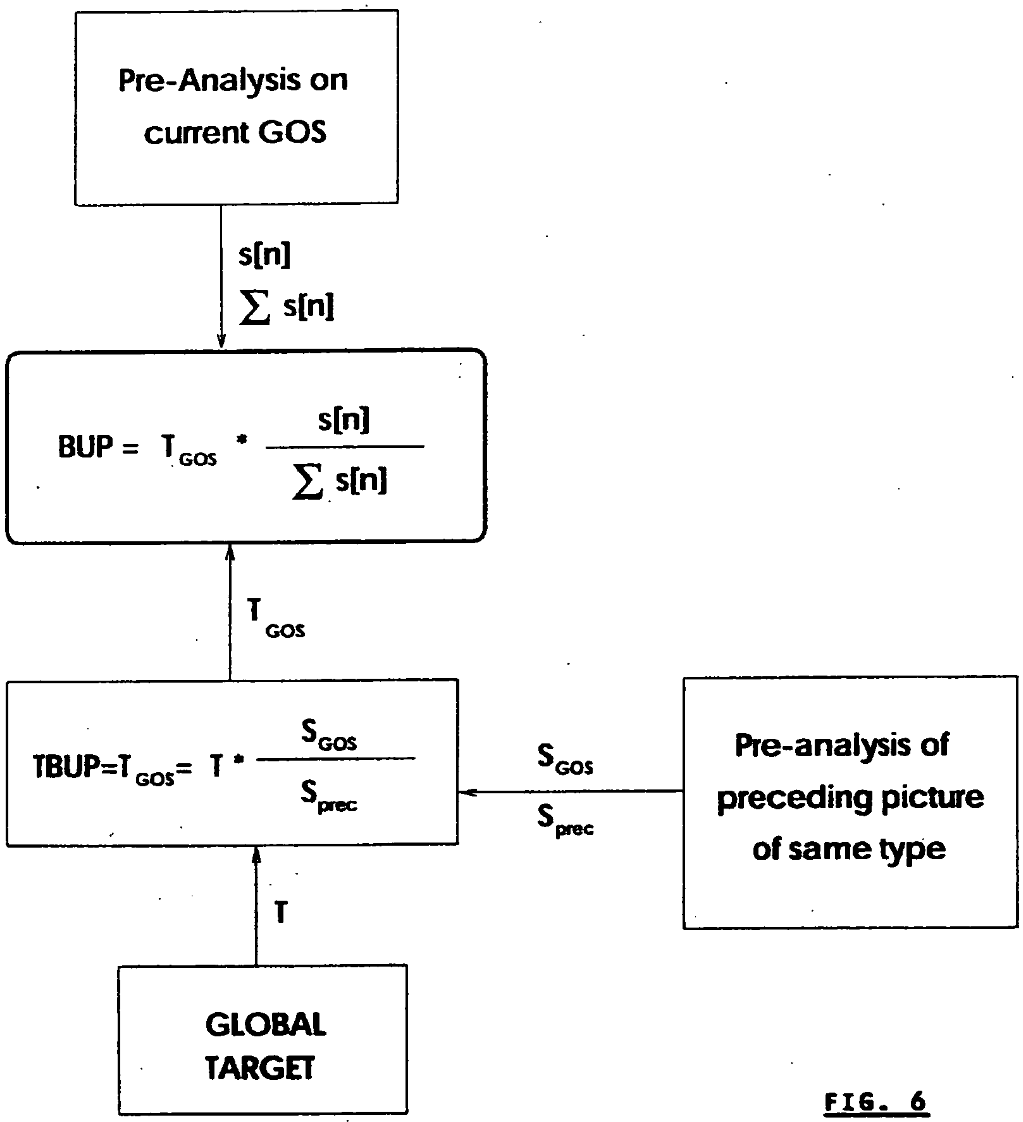
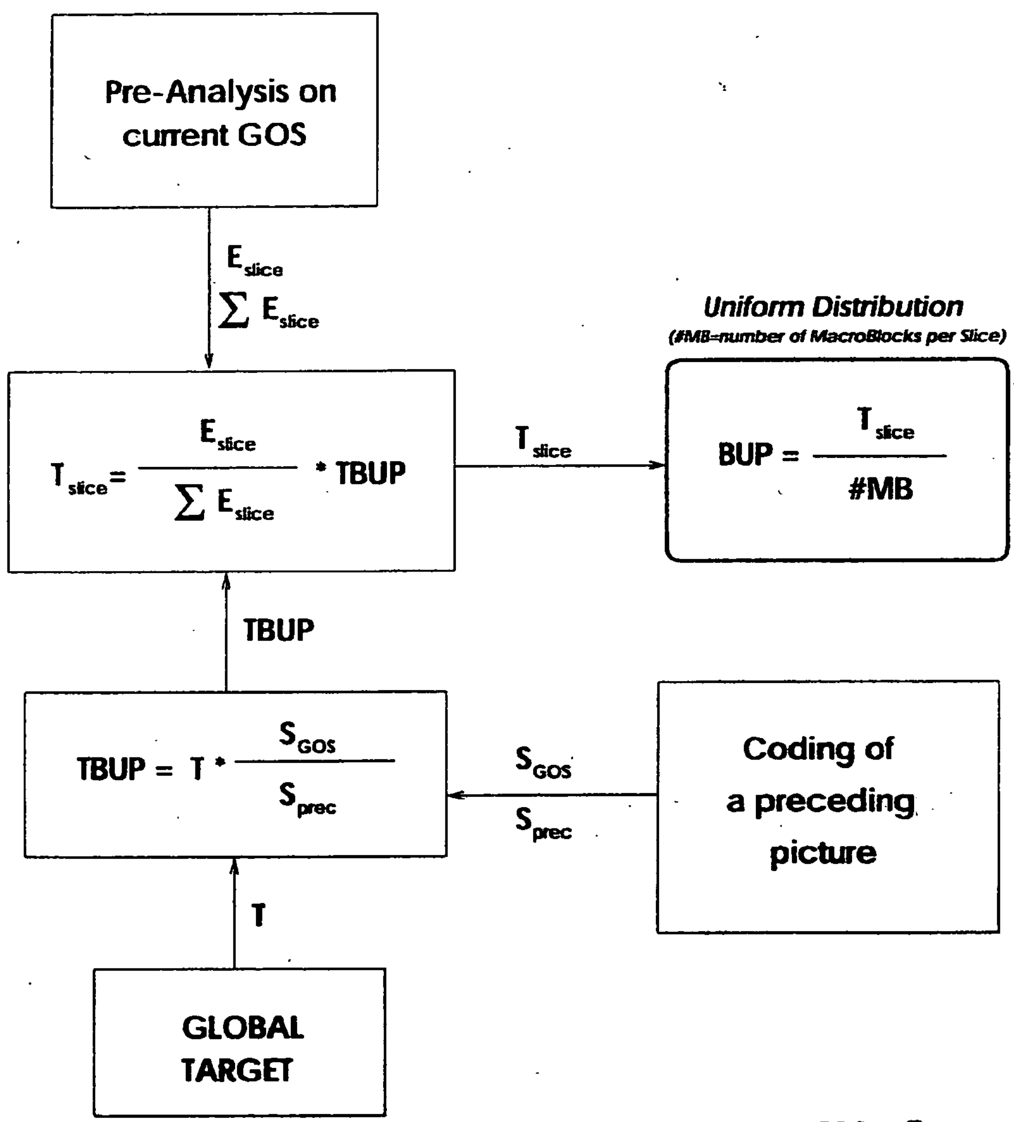

Pre-Analysis on current GOS
s[n]
∑ s[n]
BUP = T GOS * s[n] / ∑ s[n]
T GOS
TBUP=T GOS= T * S GOS / S prec
S GOS
S prec
Coding of preceding picture of same type
T
GLOBAL TARGET

FIG. 5



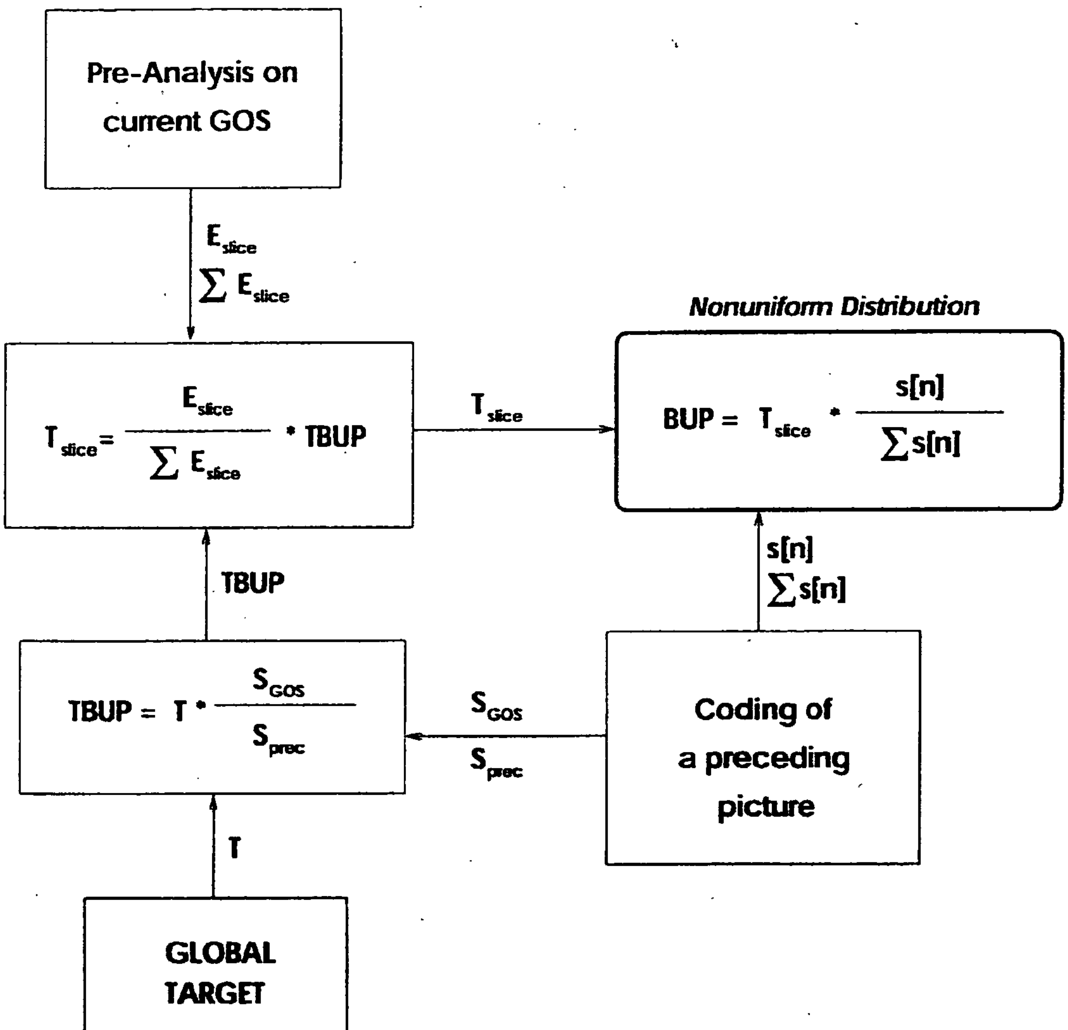
Pre-Analysis on current GOS
s[n]
$\sum$ s[n]
BUP = T_{GOS} * $\frac{s[n]}{\sum s[n]}$
T_{GOS}
TBUP=T_{GOS}= T * $\frac{S_{GOS}}{S_{prec}}$
S_{GOS}
S_{prec}
Pre-analysis of preceding picture of same type
T
GLOBAL TARGET

FIG. 6



Pre-Analysis on current GOS
E_{slice}
$\sum E_{slice}$
$T_{slice} = \frac{E_{slice}}{\sum E_{slice}} * TBUP$
T_{slice}
Uniform Distribution
(#MB=number of MacroBlocks per Slice)
$BUP = \frac{T_{slice}}{\#MB}$
TBUP
$TBUP = T * \frac{S_{GOS}}{S_{prec}}$
S_{GOS}
S_{prec}
Coding of a preceding picture
T
GLOBAL TARGET

FIG. 7



Pre-Analysis on current GOS
E_{slice}
$\sum E_{slice}$
$T_{slice} = \frac{E_{slice}}{\sum E_{slice}} * TBUP$
T_{slice}
Nonuniform Distribution
$BUP = T_{slice} * \frac{s[n]}{\sum s[n]}$
TBUP
s[n]
$\sum s[n]$
$TBUP = T * \frac{S_{GOS}}{S_{prec}}$
S_{GOS}
S_{prec}
Coding of a preceding picture
T
GLOBAL TARGET

FIG. 8

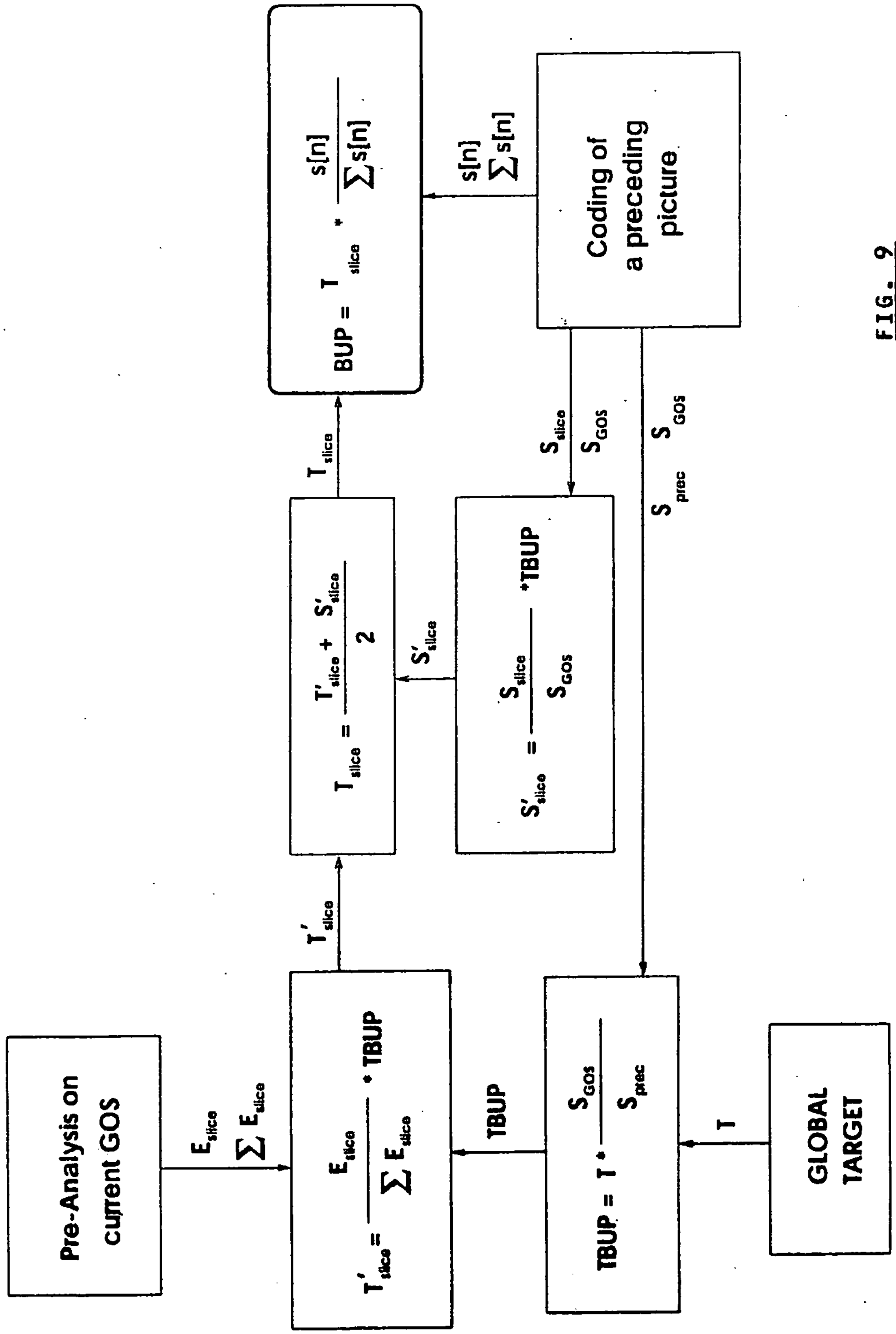

Pre-Analysis on current GOS
E slice
∑ E slice
T' slice = E slice / ∑ E slice * TBUP
T' slice
T slice = (T' slice + S' slice) / 2
T slice
BUP = T slice * s[n] / ∑s[n]
TBUP
TBUP = T * S GOS / S prec
T
GLOBAL TARGET
S' slice
S' slice = S slice / S GOS *TBUP
S slice
S GOS
S prec
S GOS
s[n]
∑s[n]
Coding of a preceding picture

FIG. 9

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 83 0599

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	BAGNI D ET AL: "Efficient intra-frame encoding and improved rate control in H.263 compatible format" , PICTURE CODING SYMPOSIUM. PCS 97, BERLIN, GERMANY, 10-12 SEPT. 1997 , ISSN 0341-0196, ITG-FACHBERICHTE, 1997, VDE-VERLAG, GERMANY, PAGE(S) 767 - 774 XP002095679 * section "3.2 Rate Control with pre-analysis" *	1-4	H04N7/50
D,Y	KEESMAN G ET AL: "BIT-RATE CONTROL FOR MPEG ENCODERS" SIGNAL PROCESSING. IMAGE COMMUNICATION, vol. 6, no. 6, 1 February 1995, pages 545-560, XP000491860 * page 548, left-hand column, line 6 - page 556, right-hand column, line 14 *	1-4	
A	US 5 721 589 A (MURATA ERI) 24 February 1998 * column 1, line 19 - column 2, line 23 *	1-4	TECHNICAL FIELDS SEARCHED (Int.Cl.6)
A	EP 0 705 041 A (SONY CORP) 3 April 1996 * page 6, line 35 - page 6, line 38 *	1-4	H04N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	5 March 1999	Fassnacht, C

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons
& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 98 83 0599

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

05-03-1999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5721589 A	24-02-1998	JP 8065677 A	08-03-1996
EP 0705041 A	03-04-1996	JP 8098185 A	12-04-1996
		US 5757968 A	26-05-1998
		US 5832121 A	03-11-1998

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82